NeuroSense: Short-Term Emotion Recognition and Understanding Based on Spiking Neural Network Modelling of Spatio-Temporal EEG Patterns

Highlights

NeuroSense: Short-Term Emotion Recognition and Understanding Based on Spiking Neural Network Modelling of Spatio-Temporal EEG Patterns

Clarence Tan, Marko Šarlija, Nikola Kasabov

- Emotion-related EEG signal segmentation strategy based on facial landmarks.
- SNN-based framework for subject-independent short-term emotion recognition from EEG.
- SNN hyperparameter optimisation related to spike encoding and output representation.
- The emerged SNN connectivity helps in understanding the neural mechanisms of emotion.

Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 00 (2020) 1–17
Neuro-computing

NeuroSense: Short-Term Emotion Recognition and Understanding Based on Spiking Neural Network Modelling of Spatio-Temporal EEG Patterns

Clarence Tan[a,*], Marko Šarlija[b], Nikola Kasabov[a]

[a] *Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Private Bag 92006, Auckland 1010, New Zealand*
[b] *Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, Zagreb 10000, Croatia*

Abstract

Emotion recognition still poses a challenge lying at the core of the rapidly growing area of affective computing and is crucial for establishing a successful human-computer interaction. Identification and understanding of emotions are achieved through various measures, such as subjective self-reports, face-tracking, voice analysis, gaze-tracking, as well as the analysis of autonomic and central neurophysiological measurements. Current approaches to emotion recognition based on electroencephalography (EEG) mostly rely on various handcrafted features extracted over relatively long time windows of EEG during participants exposure to appropriate affective stimuli. In this paper, we present a short-term emotion recognition framework based on spiking neural network (SNN) modelling of Spatio-temporal EEG patterns. Our method relies on EEG signal segmentation based on detection of short-term changes in facial landmarks, and as such includes no computation of handcrafted EEG features. Differences between participants' EEG properties are taken into account via subject-dependent spike encoding in the formulated subject-independent emotion recognition task. We test our methods on the publicly available DEAP and MAHNOB-HCI databases due to the availability of both EEG and frontal face video data. Through an exhaustive hyperparameter optimisation strategy, we show that the proposed SNN-based representation of EEG spiking patterns provides valuable information for short- term emotion recognition. The obtained accuracies are 78.97% and 79.39% in arousal classification, and 67.76% and 72.12% in valence classification, on the DEAP and MAHNOB-HCI datasets, respectively. Furthermore, through the application of a brain-inspired SNN model, this study provides novel insight and helps in the understanding of the neural mechanisms involved in emotional processing in the context of audiovisual stimuli, such as affective videos. The presented results encourage the use of the proposed EEG processing methodology as a complement to existing features and methods commonly used for EEG-based emotion recognition, especially for short-term arousal recognition.

Keywords: Spiking Neural Networks, Emotion Recognition, Affective Computing, EEG, Event Detection

1. Introduction

The interaction between humans and computational devices is becoming more and more common with the advent of personal digital devices, wearable systems, and other technological interventions. The field of affective computing (AC) combines computer science and emotion research to enable computational systems to identify the emotional states of users (affect recognition) and to generate responses that humans are likely to perceive as emotional (affect generation). It has also been argued that over time, it may be possible for systems to actually "feel" emotions [1].

*Corresponding author
Email addresses: cltan@aut.ac.nz (Clarence Tan), marko.sarlija@fer.hr (Marko Šarlija), nkasabov@aut.ac.nz (Nikola Kasabov)

The introductory work by Picard [1]. was followed by much research at the intersection of diverse fields such as neuroscience, ethics, psychology, and engineering, among others. To this day, the work on AC has resulted in the improvement of systems that are capable of interpreting, identifying, and responding to the emotional states of users. For this purpose, affective computing makes use of various multimodal inputs such as facial images, i.e. facial expression, voice data, biometric data, e.g. physiological changes, and body language of the user. Computational models termed as "affect models" are then employed to make sense of these input parameters and identify the emotional state of the user [2]. The significance of AC is on the rise, given the increased degree of human-computer interactions. According to a recent study[1], 95% of the American population now owns a cell phone of some kind among which 77% use a smartphone. To compare with, the percentage of smartphone users was 35% in 2011. Instead of the one-sided interactions that humans normally expect from machines, utilising AC can make these systems respond in more effective ways, making the whole technology experience more satisfactory to the user [3]. Devices often come with built-in sensors that collect user data. The key challenge, lying at the core of AC is recognising the emotional state of a person based on the available data.

Nowadays, AC is finding applications in various fields ranging from gaming [4], education and e-learning [5, 6] to medicine [7, 8], wearable devices [9], robotics [10], etc. For example, research shows that effective computing systems may aid in the diagnosis of seemingly hidden and unobservable medical conditions such as depression and chronic pain [8]. Similarly, affective systems can provide more empathetic, personalised feedback to students, making online learning more efficient [11]. As briefly stated earlier, in order to sense affective states, four types of inputs are typically used: 1) facial expression recognition [12, 13, 14], 2) voice recognition [15, 16, 17], 3) gesture recognition [18, 19] and 4) biometrics [20, 21, 22]. Various instruments are used to gather the aforementioned types of data, like cameras, microphones, sensors and biometric devices (e.g. heart rate, blood pressure, skin conductance or electroencephalography (EEG) measurements). The vast majority of AC research if focused on detailed analysis of the collected data, in order to determine the emotional state of the user, where various machine learn- ing techniques have been employed [17, 23, 24]. Affect models are usually trained on large datasets of relevant data so that they can then be employed for emotion recognition and affect generation.

Various improvements in model architectures, feature selection methodology [20] and deep-learning-based data representations have led to increases in emotion classification accuracies in the last years. However, exploration of novel approaches and concepts in the analysis of human affect could add new information and thus complement traditionally used approaches and features. As stated above, EEG is one of the well-established modalities in affective research [12, 13, 25]. On the other hand, spiking neural networks (SNNs) have recently proven to be successful in modelling, recognition and understanding of EEG Spatio-temporal data in a wide array of domains [26], as described in section 2. Based on the importance of EEG in AC [25], as well as numerous EEG-based applications using SNNs (described in section 2), in this paper, we adapt and apply the concept of evolving spiking neural networks (eSNN) [27] and propose an SNN classification framework for EEG-based short-term emotion recognition.

Main contributions of our paper are:

- EEG signal segmentation strategy based on detection of changes in facial landmarks for short-term emotion recognition (section 3.2).
- SNN-based framework for subject-independent short-term emotion recognition (section 4).
- Hyperparameter optimisation strategy for spike encoding and the dynamic evolving SNN (deSNN) data representation (subsections 4.1.1 and 4.2.2).
- Comparison of emotion classification accuracies obtained by simple EEG spike-based features vs complex SNN-based representation of EEG spiking patterns (section 5).
- Novel SNN-based insights related to the neural mechanisms involved in short-term emotional processing of affective videos.

[1] https://www.pewresearch.org/internet/fact-sheet/mobile/

2. Spiking neural networks

Human brains encode information via discrete events known as action potentials or spikes, following an all-or-none principle, where a neuron fires a spike once the accumulated potential reaches a certain threshold, else it remains silent. Due to this binary nature of information representation, the human brain still outperforms the traditional artificial neural networks (ANNs) in terms of both energy and efficiency [28, 29]. Compared to the traditional ANNs, SNNs utilise a more biologically plausible model of neurons [30], thus bridging further the gap between neuroscience and learning algorithms. SNNs have shown the ability to integrate information encoded in time, phase, frequency, as well as handle large volumes of data in an adaptive and self-organised manner [31], making them particularly suitable for solving online Spatio-temporal pattern recognition problems. SNNs have been shown to be computationally more efficient than ANNs, both theoretically [32, 33] and in several real-world applications [34]. SNNs have over the past years proven to be successful in several real-world learning tasks such as unsupervised classification of non-globular clusters [35], image segmentation and edge detection [36], as well as in various tasks related to modelling, recognition and understanding of EEG Spatio-temporal data [37, 38, 39], such as Alzheimer's disease classification [40], epilepsy and epileptic seizure detection [41], predicting human behaviour during decision making [42], detection of limb movement execution and intention for brain-machine interface (BMI) applications [43], classification of activities of daily living [44], modelling of peri-perceptual brain processes [45], distinguishing brain states associated with depression and responsiveness to Mindfulness Training [46], etc.

The evolving SNN, i.e. eSNN, is a class of SNN that utilises rank order learning [47] and was first proposed in [48]. In addition to the open evolving structure which facilitates the addition of new variables and neuronal connections [49], eSNNs have the advantage of fast learning from large amounts of data and can interact with other systems actively. eSNNs also allow the integration of various learning rules such as supervised learning, unsupervised learning, fuzzy rule insertion and extraction, to mention a few, and are self-evaluating in terms of system performance. These aforementioned properties constitute the evolving connectionist systems (ECOS) principles on which the eSNN is based [50]. In the rank-order learning scheme, the synaptic weights are adjusted only once, making it not very efficient for Spatio-temporal data, where there may be a need to adjust synaptic weights based on the spikes arriving on a given synapse over time. To overcome this disadvantage, an extension of eSNN known as dynamic eSNN (deSNN) was introduced in [31] that combines rank-order learning with temporal learning rules such as spike-timing-dependent plasticity (STDP), which allows dy- namic adjustment of the synaptic weights (more details in subsection 4.2.2). However, both eSNN and deSNN do not encapsulate the structural information of the brain in terms of neuronal locations and their connectivity, which may be crucial for modelling of Spatio-temporal brain data (STBD), such as EEG. The NeuCube architecture, first proposed in [51], aims at building an eSNN that incorporates structural as well as functional aspects of the brain along with utilising the unsupervised STDP learning algorithm. Below we give a brief introduction to the NeuCube architecture.

Traditional supervised learning methods such as support vector machines (SVM) or multilayer perceptron neural networks (MLP) typically deal with the spatial or temporal aspects of brain data, but cannot handle the dynamic interaction between these processes [38]. Furthermore, such models cannot incorporate any prior structural knowledge of the brain in an unsupervised manner, or handle multimodal brain data, e.g. EEG, functional magnetic resonance imaging (fMRI), diffusion tensor imaging (DTI), positron emission tomography (PET). NeuCube is a specific implementation of an eSNN, initially proposed to handle pattern recognition problems related to STBD, but has further been developed and modified in order to handle various other types of Spatio-temporal data such as audiovisual data, climate data, seismic data and ecological data [26, 52]. The emotion recognition methodology proposed in this paper is based on the NeuCube framework, which has been adapted and further developed for this specific task, particularly in the directions of subject-specific spike encoding and hyperparameter optimisation (see section 4).

3. Data preparation

3.1. Datasets

Due to the high interest in EEG analysis for emotion recognition, several publicly available multimodal databases have been established to this day, like DEAP [12], MAHNOB-HCI [13], SEED [53] or DREAMER [54], all of which include EEG. In this paper, we use DEAP and MAHNOB-HCI databases due to their similarity and availability of frontal face video data, which is needed for our event-detection-based signal segmentation.

DEAP is a widely used dataset for multimodal emotion analysis, consisting of 32 subjects [13]. 32-channel EEG and peripheral physiological signals, namely the galvanic skin response (GSR), electrooculogram (EOG), electromyogram (EMG), respiration, plethysmograph, electrocardiogram (ECG) and body temperature were recorded during subjects' exposure to 40 one- minute long affective music videos. After watching each video, each subject rated their emotional experience in five dimensions: valence, arousal, dominance, liking and familiarity. The rating values were on a continuous scale of 1-9, except for familiarity, which was rated on a discrete scale of 1-5. For 22 out of the 32 subjects, the frontal video was recorded, so we used data from those 22 subjects only.

MAHNOB-HCI is a multimodal dataset consisting of 27 subjects which participated in two experiments [13]. In the first experiment, similarly to DEAP, each participant watched 20 emotional videos which were between 34.9 and 117 seconds long. Recorded signals included 32-channel EEG, peripheral physiology (ECG, GSR, respiration, skin temperature), eye gaze data, audio data and video from 6 cameras recording facial expressions and head pose. After watching each video, the participants gave an emotional label/tag to the video, as well as rated their emotional experience in arousal, valence, dominance and predictability on a 1-9 scale. The second experiment was related to implicit tagging and was not used in this work.

3.2. EEG segmentation based on the analysis of facial landmarks

In this section, we describe the signal segmentation procedure, which was used to detect moments of participants' most intensive emotional engagement, based on detecting events in the time-varying facial landmarks. This step is important due to several reasons:

- The differences in utterance lengths vary between DEAP and MAHNOB- HCI, as well as within the MAHNOB-HCI database itself. With an event-based approach, we make sure all samples which are to be used for later emotion recognition are of the same length.
- The self-assessed emotional experience of the participant is not uniformly distributed across the entire duration of an emotional video but is more likely a result of one or more emotionally intensive events of shorter duration. We aim to capture the occurrences of such events by the analysis of facial landmarks.

The facial-video-based event detection algorithm comprises the following steps, which were taken in both datasets:

1. In each frame of each video, the participant's face was tracked. We have employed an implementation[2] of the Viola-Jones algorithm [55] to detect participant's faces, noses, eyes, mouth, etc., in the first frame of the video. This step outputs a region of interest (ROI).
2. ROI from the previous step is then tracked from frame to frame based on detection and tracking of specific features by using a minimum eigenvalue feature detection algorithm developed by Shi and Tomasi [56] and a Kanade–Lucas–Tomasi (KLT) feature tracking algorithm [57].
3. Each detected and tracked ROI from the previous step was used as input to a facial landmarks detection algorithm[3] [58]. Facial landmarks were 68 specific points on the face, such as mouth corners, eyebrow lines, eye lines, etc., as shown in Figure 1.
4. Based on the detected and tracked 68 facial landmarks, we compute an array of 20 specific geometrical features per frame, related to eyebrow, eye and lip positioning and shape, as described in [13].
5. The energy of time-varying geometrical facial features (see Figure 2, top) is calculated disregarding the 10 features based on eye landmarks (features f5 to f14 in [13]), as the eye-based features are very sensitive to blinking. The obtained energy signal is shown in Figure 2, bottom.
6. A simple liner-slope-based match filter for detecting increases in the facial features energy signal is applied, resulting in a detection signal (see Figure 2). A maximum in the detection signal marks the beginning of a one-second-long event. Figure 3 shows the face of the participant at the beginning and ending of the first event from Figure 2.

[2]The *vision.CascadeObjectDetector* object from Matlab's Computer Vision Toolbox.
[3]The pretrained DLIB model for facial landmarks detection: http://dlib.net/files/shape_predictor_6_face_landmarks.dat.bz2

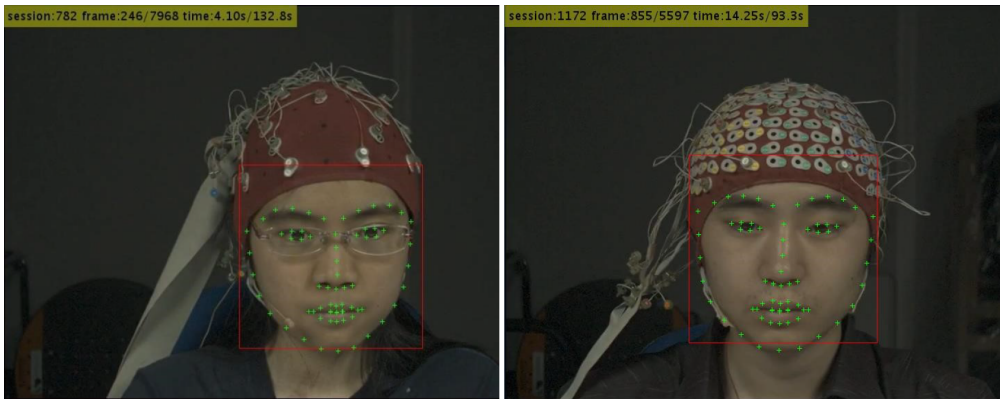

Figure 1: Two examples of facial landmarks detection in the MAHNOB-HCI dataset.

7. EEG signal segmentation is finally performed based on the detected events from the previous step. EEG signals were preprocessed using the TEAP toolbox [59], and all 32 available channels are used.

The described procedure resulted with a total of 224 samples for arousal classification and 224 samples for valence classification from the DEAP dataset, and a total of 162 samples for arousal classification and 208 samples for valence classification from the MAHNOB-HCI dataset. Data from participants with less than 5 samples per participant were excluded from the analysis, resulting with a total of 214 samples for arousal classification (125 labelled positively) and 214 samples for valence classification (112 labelled positively) from the DEAP dataset, and a total of 131 samples for arousal classification (47 labelled positively) and 191 samples for valence classification (94 labelled positively) from the MAHNOB-HCI dataset.

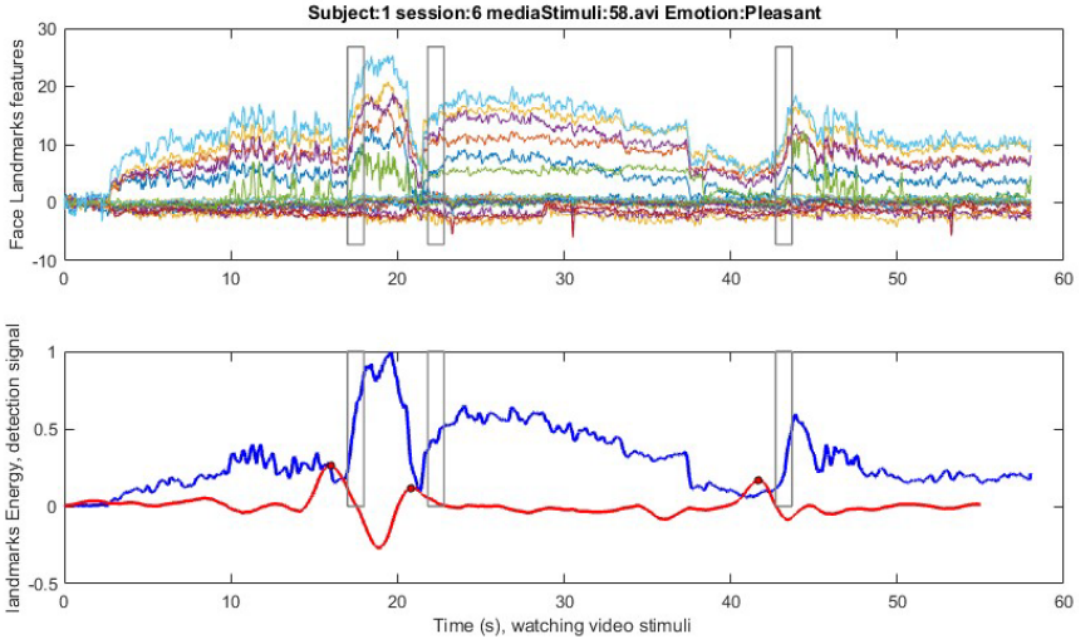

Figure 2: Detection of increases in facial activity. The top axes are showing the trajectories of facial features during a participant's exposure to a pleasant video. The bottom axes are showing the corresponding facial features energy signal (blue) and the detection signal (red). In this example, three events were detected.

Figure 3: An example of a change in the participant's facial expression from the beginning to the ending of a detected emotional event.

The number of extracted samples, as well as the class distributions, are not equal for valence and arousal due to our labelling strategy. For arousal, we used utterances which were labelled as either *calm* or *excited/activated*, thus excluding *medium arousal*. Accordingly, for valence, we also selected 2 classes (*pleasant* and *unpleasant*), thus excluding utterances labelled as *neutral valence*. This resulted with cases for which a participant rated his experience as, e.g. *pleasant* and *medium arousal*, in which case the data is processed for valence analysis and not processed for arousal analysis.

4. Emotion recognition methodology

In the following section, we describe our emotion recognition methodology based on SNN modelling of Spatio-temporal EEG spike patterns. Accord- ing to recent reviews of various studies on emotion recognition from EGG [25, 60], static features extracted on specific slices of EEG data dominate the research landscape. These are most commonly related to spectral features, like changes in power over the theta, alpha, beta and gamma frequency bands [12, 61, 62] and spectral power asymmetry measures in pairs of symmetrical electrodes [13, 62]. Besides the frequency domain features, which are predominant, various time domain and nonlinear features have been investigated as well [63], such as Hjorth features [64], fractal dimensions, entropy features etc. Regardless of the feature extraction methodology, traditional approaches usually result in a static feature vector, thus usually well capturing the spatial aspect (emerging from the electrode positions), but neglecting the dynamical Spatio-temporal nature of EEG patterns. Time-frequency domain features are currently the only tool used to capture these dynamical changes in EEG [65]. With our work, we aim to leverage the Spatio-temporal nature of EEG in emotion recognition by investigating the concept of spike encoding and SNN modelling of EEG. Therefore, we focus only on the comparison of the predictive power of simple spike-based features and complex SNN based Spatio-temporal spike-patterns in the task of EEG based emotion recognition. Accordingly, the possibility of enhancing the predictive power by the integration of EEG-spike-based features with traditional static features described earlier exceeds the scope of our work.

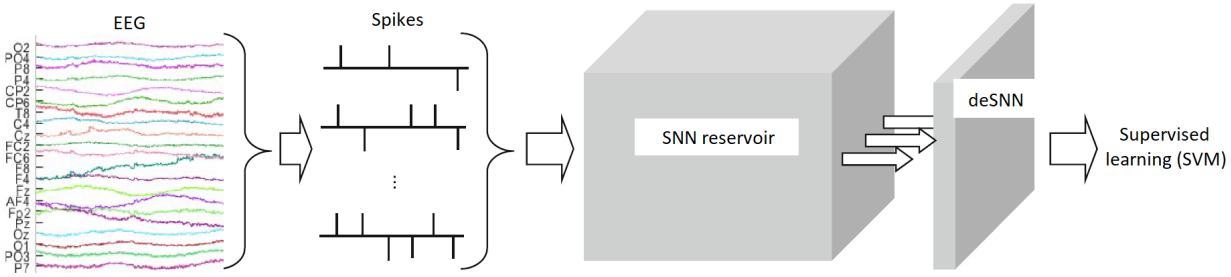

Figure 4: Illustration of the proposed SNN computational architecture for EEG-based emotion recognition.

The proposed SNN-based framework for subject-independent emotion recognition includes the following processing steps (illustrated in Figure 4):

1. Subject-specific spike encoding of short-term EEG recordings obtained by the segmentation strategy described in section 3.2.
2. Unsupervised learning of a brain-like 3-D SNN reservoir (SNNr) module, based on STDP learning.
3. deSNN representation of Spatio-temporal spiking patterns obtained from the trained SNNr, given a specific input spike sequence. The obtained representation forms a static vector which is suitable as input to any traditional supervised learning algorithm (e.g. SVM, ANN, kNN, etc.).
4. Feature selection and supervised learning for emotion recognition from the representation obtained in the previous step.

4.1. Spike encoding

In an SNN-based architecture, information is processed in the form of binary spiking events. Accordingly, all continuous variables first need to be encoded into spike trains, shown as the first step in Figure 4. In this paper, by spike encoding, we namely focus on temporal spike encoding methods where spike timings usually mark changes in the signal value over time [66]. This approach is driven by a biologically plausible view that precise relative spike timing encodes information [67]. Most of the commonly used encoding algorithms [52], such as threshold-based representation (TBR), step-forward (SF) encoding, moving-window (MW) encoding, or the Bens Spiker Algorithm (BSA) [68], rely on tracking the temporal changes in the signal, which then represent the exact timing of spikes. Such algorithms produce a bipolar spike sequence, where positive changes in signal value (increases) result with positive spikes, and negative changes in signal value (decreases) result with negative spikes.

In this paper, we transform the EEG data into spike trains using the relatively simple version of TBR [69, 70]. The method is based on thresholding the rate of change of an input variable over time and is suitable when the input data is a stream, which is the case with EEG. The algorithm is based on the variable threshold value that is calculated for each of the 32 input data channels. The variable threshold array is calculated for each channel, as the signal dynamics and value ranges can vary between the input channels. For each of the input channels, the variable threshold is calculated based on one scalar input parameter (α_{TR}) in the following way:

$$VT(k) = \frac{1}{N}\sum_{i=1}^{N}\left(\mu + \sigma \cdot \alpha_{TR}\right) \tag{1}$$

where N is the number of samples, T is the signal length (number of time points per data sample), and k goes from 1 to the number of channels $N_{\text{input}} = 32$. α_{TR} is the spike threshold parameter, μ is the sample mean rate of change in the signal, σ is the sample standard deviation of the rate of change in the signal, and $\boldsymbol{VT}$ is the resulting variable threshold array. Equation (1) represents the threshold value for the k-th input channel. At each time point where the k-th input channel signal difference (rate of change) exceeds the corresponding variable threshold, a positive spike is generated. Accordingly, inhibiting, i.e. negative, spikes are generated when the rate of change exceeds the variable threshold in the negative, i.e. decreasing, direction. Algorithm 1 sums up the spike encoding procedure in a compact algorithmic form.

In affective applications, such as the subject-independent emotion recognition, the encoding algorithm should ideally provide similar spike trains for similar emotional states, regardless of the high inter-subject variability in the collected EEG signal properties. Therefore, the encoding procedure described in algorithm 1 is utilised in a per-subject fashion, with μ and σ being estimated separately for each subject.

4.1.1. Optimisation of the spike encoding method

Spike encoding is the first link in the SNN processing chain. Choosing the appropriate spike encoding method, with the optimal hyperparameters, in our case α_{TR}, is extremely important, in order to retain the task-relevant information. The spike encoding can therefore be seen both as a primary information compression as well as a noise

Algorithm 1: Spike encoding: $f_{encode} : \mathbb{R}^{T \times N_{input}} \rightarrow \{-1, 0, 1\}^{T \times N_{input}}$

Require: $X_{in} \in \mathbb{R}^{T \times N_{input}}, \{hyperparameters := \alpha_{TR}\}$
Ensure: $X_{out} \in \{-1, 0, 1\}^{T \times N_{input}}$
$N \leftarrow \#(X_{in})$ {number of data samples in the dataset}
for $k = 1$ to N_{input} **do**
  $VT_k \leftarrow 0$
  for $i = 1$ to N **do**
    $x \leftarrow$ channel k of the i-th sample in X_{in}
    $x' \leftarrow |\delta x|$
    $\mu \leftarrow mean(x')$
    $\sigma \leftarrow st.dev.(x')$
    $VT_k \leftarrow VT_k + (\mu + \sigma \cdot \alpha_{TR})$
  end for
  $VT_k \leftarrow \frac{VT_k}{N}$
end for
for $k = 1$ to N_{input} **do**
  for $i = 1$ to N **do**
    $x \leftarrow$ channel k of the i-th sample in X_{in}
    $x' \leftarrow \delta x$
    $x_{out} \leftarrow 0^{T \times 1}$
    for $j = 2$ to T **do**
      if $x'_j > VT_k$ **then**
        $x_{out(j)} \leftarrow 1$
      else if $x'_j < -VT_k$ **then**
        $x_{out(j)} \leftarrow -1$
      end if
    end for
    store spike train x_{out} in X_{out} for channel k, sample i
  end for
end for

elimination step. Inadequate spike encoding algorithm can result with either loss of useful information on one end (high α_{TR}), or high levels of information noise on the other end (low α_{TR}), as shown in Figure 5. The problem of selection and optimisation of temporal spike encoding algorithms for SNNs has been most recently tackled in [66]. An exhaustive approach should evaluate the effectiveness of the encoding within the context of the entire SNN processing and classification framework, i.e. the encoding algorithm is optimised according to the output of the whole SNN system, e.g. classification accuracy [71]. The computational cost of such an approach would be extremely high, due to the need for simultaneous optimisation of all SNN system hyperparameters, thus making the approach infeasible. Another approach is to try and optimise the encoding step by itself, via application of the corresponding decoding algorithm and trying to minimise some error metric between the original and reconstructed signal, as in [66]. This approach neglects the classification task context, thus lacking final validation in terms of the obtained classification performance.

In this paper, we propose a compromise approach: we optimise the encoding method by itself, due to the already stated computational infeasibility of taking into account the rest of the SNN system, while we still take into account the specific classification problem, i.e. binary emotion classification in terms of valence and arousal. In order to avoid evaluating the performance of the entire SNN system, we calculate an array of simple spike-based features based on statistical descriptives and the widely used rate coding scheme [72]. For each of the 32 EEG channels, encoded into spike trains as described in the previous subsection, we calculate the following 6 features:

- firing rate of positive spikes,
- firing rate of negative spikes,
- median timing difference between consecutive positive spikes,

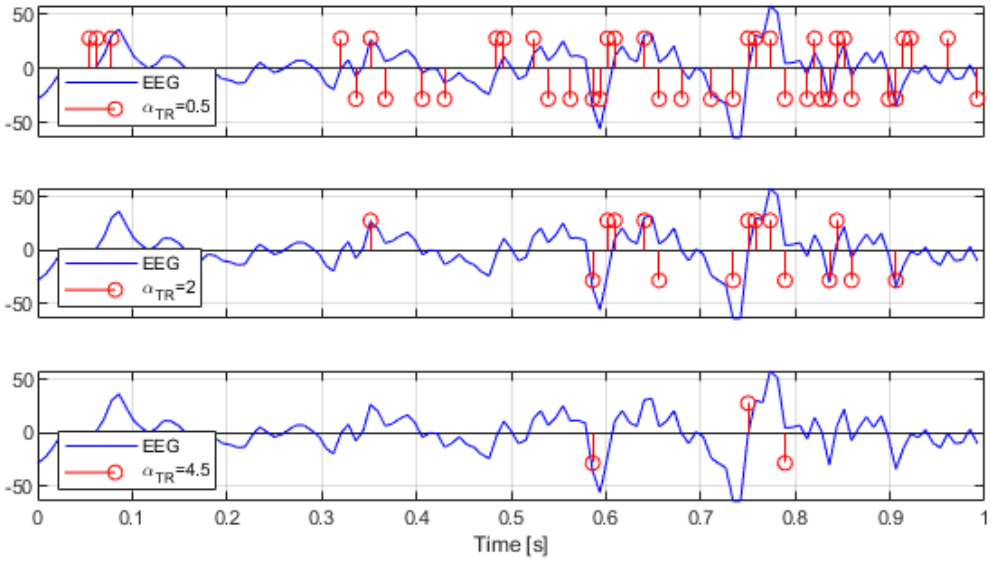

Figure 5: Spike trains resulting from different AER threshold parameter (α_{TR}) values. Lower α_{TR} yields a dense spike train with a tendency to encode changes which are most likely on the level of noise, while high α_{TR} value yields a sparse spike train with only major changes in signal value being encoded as spike events.

- median timing difference between consecutive negative spikes,
- interquartile range of timing differences between consecutive positive spikes,
- interquartile range of timing differences between consecutive negative spikes.

We use a simple definition of firing rate as the temporal average, i.e. the number of spikes divided by the corresponding time interval duration (in our case 1 second for all data samples). Each sample is therefore described by $32 \cdot 6 = 192$ simple rate-coding-based and statistical features. Datasets are formed for an array of different α_{TR} values (ranging from 0.5 to 5, with a step of 0.5), standardised, and then tested for their task-relevant discriminative power. The used performance metric is the cross-validated classification error obtained with an optimised Inf-FS feature selection algorithm [73, 74] and an RBF-SVM classification algorithm. This metric employs a simple and generic feature selection and machine learning approach in order to estimate the maximum cross-validated nonlinear class separation that can be obtained by the calculated features. For each α_{TR} value, a non-convex Bayesian optimisation approach was employed in order to find the values of the remaining hyperparameter values:

- α_{FS}: Inf-FS parameter representing the trade-off between feature dispersion and feature correlation, which are the basis of the Inf-FS feature ranking algorithm (range set to $[0, 1]$),
- n_{FS}: number of selected features, based on the obtained feature ranking (range set to $[1, 50]$, integer grid),
- C: RBF-SVM parameter (range set to $[10^{-3}, 10^{3}]$, logarithmic grid),
- σ: RBF-SVM parameter (range set to $[10^{-3}, 10^{3}]$, logarithmic grid),

which maximize the leave-one-subject-out (LOSO) cross-validation accuracy. Matlab function *bayesopt* was used, with 500 objective evaluations and a 0.5 exploration ratio. Algorithm 2 sums up the spike encoding optimisation procedure in a compact algorithmic form.

Algorithm 2 has been applied to the tasks of arousal and valence classification, on both DEAP and MAHNOB-HCI, and the results of the spike encoding optimisation are shown in Figure 6. Threshold value $\alpha_{TR} = 1.5$ yields the most informative spikes for both arousal and valence classification. Such a result can be interpreted in the context of the behaviour of the encoding algorithm illustrated in Figure 5.

Algorithm 2: Optimisation of the spike encoding threshold parameter α_{TR}

Require: $X_{in} \in \mathbb{R}^{T \times N_{input}}$, $y \in \mathbb{R}^{N_{samples} \times 1}$
Ensure: $\alpha_{TR,opt.}$
1: $\alpha_{TR,range} \leftarrow \{0.5k, k \in \{1, 2, ..., 10\}\}$ {search range for α_{TR}}
2: $N_{search} \leftarrow \#(\alpha_{TR,range})$ {from the above: $\#(\alpha_{TR,range}) = 10$}
3: $CV_{loss} \leftarrow 0^{N_{search} \times 1}$ {initialise objective array}
4: $\alpha_{FS,range} \leftarrow [0, 1]$) {search range for α_{FS}}
5: $n_{FS,range} \leftarrow [1, 50]$) {search range for n_{FS}, integer}
6: $C_{range} \leftarrow [10^{-3}, 10^{3}]$) {search range for C, logarithmic grid}
7: $\sigma_{range} \leftarrow [10^{-3}, 10^{3}]$) {search range for σ, logarithmic grid}
8: **for** $i = 1$ to N_{search} **do**
9: $X \leftarrow f_{encode}(X_{in}, \alpha_{TR,range}(i))$ {spike encoding, see algorithm 1}
10: calculate $X_{features} \in \mathbb{R}^{N_{samples} \times 192}$ based on X {features described in section 4.1.1}
11: $CV_{loss}(i) \leftarrow Bayesopt(X_{features}, y, \alpha_{FS,range}, n_{FS,range}, C_{range}, \sigma_{range})$ {cross-validated classification error, other hyperparameters described in 4.1.1}
12: **end for**
13: $i_{min} \leftarrow$ index of smallest element in CV_{loss}
14: $\alpha_{TR,opt.} \leftarrow \alpha_{TR,range}(i_{min})$

4.2. Spiking neural network processing

4.2.1. 3D SNNr module

Once the EEG data has been encoded for each subject, the obtained sparse spike trains are used to train a 3D SNNr with $N_r = 1471$ leaky integrate and fire (LIF) model neurons. Each neuron has a predefined 3D spatial coordinate, according to the Talairach template coordinates from [75], resulting with a brain-like shape. Accordingly, input spikes are passed over to the SNNr at the neuron locations corresponding to the mapping of 32 EEG channels, as shown in Figure 7.

The connections between the neurons in the SNNr are initialised using the small-world connectivity (SWC) approach [37], where a radius is defined as a parameter for connecting neurons within this radius, i.e. small-world radius (SWR). This results in an SNNr of sparsely connected neurons. After initialisation, the connection weights

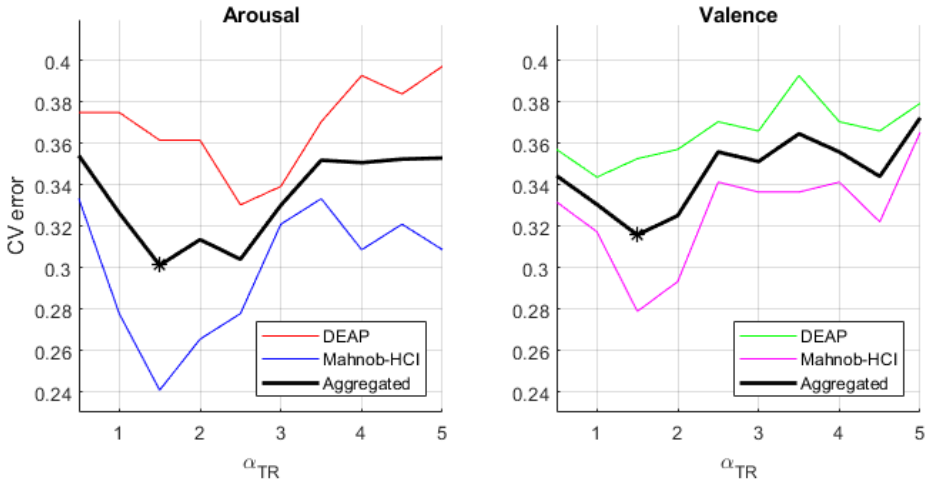

Figure 6: Class separation performance metrics for different α_{TR} values, using both datasets: arousal classification (left) and valence classification (right). Optimal α_{TR} value is indicated by an asterisk.

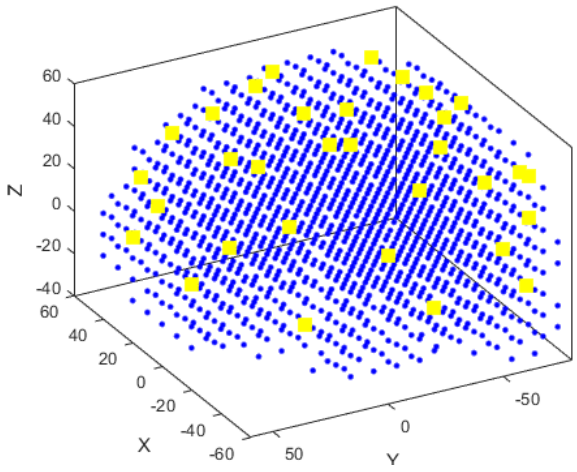

Figure 7: 3D SNNr structure, according to the Talairach template coordinates. Yellow neurons are considered input neurons, and correspond to the mapping of 32 EEG channels.

$W_{(ij)}$ between the pairs of connected neurons (ij) are determined based on the following expression:

$$W_{(ij)} = \text{sgn}\Big(\text{rand}(1) - 0.2\Big) \cdot \text{rand}(1) \cdot \frac{1}{L_{\text{dist}(ij)}}, \tag{2}$$

Where rand(1) generates a pseudorandom from a uniform distribution on the open interval (0, 1), and $L_{\text{dist}(ij)}$ represents the distance between neurons i and j. The equation above results with an expectation of 80% positive weights and 20% negative weights (in matrix $\boldsymbol{W}$).

The STDP learning rule [76] is applied that allows the SNNr to adapt the connection weights based on the Spatio-temporal associations between the input-driven spikes. The used STDP algorithm relies on the following hyperparameters:

- D (potential leak rate): the rate of potential passive degradation through inactivity,
- R (refractory time): determining a period of resting between spikes,
- η (STDP rate): learning rate, used for weight updating,
- β (firing threshold): potential threshold for generating a spike,
- N_{iter}: number of training iterations,

and with the specific steps described in Algorithm 3.

4.2.2. *deSNN representation*

As an output representation of the SNNr-based spike sequences we used the deSNN algorithm [31]. The method combines the simple rank-order (RO) learning rule [47] and the STDP-based activation of the neurons in the previously trained SNNr (as described in Algorithm 3). In the previous section, we have described how the spike trains can be used to train the SNNr in an unsupervised manner. However, the SNNr now operates as an activation module, meaning that the new input spike trains are propagated through the trained SNNr, generating a complex Spatio-temporal neuron activation pattern. This pattern is used to generate (evolve) the output neurons, i.e. the deSNN representation. The algorithm hyperparameters are:

- α_{deSNN}: main deSNN hyperparameter that determines the output value based on the first spike occurrence of the corresponding SNNr neuron,
- d(drift): used for output update on the subsequent spikes of the corresponding SNNr neurons.

Algorithm 3: Unsupervised SNNr weight learning: f_{STDP}

Require: $C_{init} \in \{0,1\}^{N_r \times N_r}$, $W_{init} \in \mathbb{R}^{N_r \times N_r}$, $S_{in} \in \{-1,0,1\}^{T \times N_{input} \times N_{samples}}$ $\{hyperparameters := D, R, \eta, \beta, N_{iter}\}$
Ensure: $W_{out} : \mathbb{R}^{N_r \times N_r}$

$\chi \leftarrow [1, 2, ..., N_r]$ {all neuron indices}
$P_k \leftarrow 0, \forall k \in \chi$ {initialize neuron potentials}
$R_k \leftarrow 0, \forall k \in \chi$ {initialize neuron refractory time counters}
find $inputneuron$ indices $\iota \subset \chi$
for $n_{iter} = 1$ to N_{iter} **do**
  $\eta' \leftarrow \frac{\eta}{\sqrt{n_{iter}}}$
  for $i = 1$ to $N_{samples}$ **do**
    $s \leftarrow T \times N_{input}$ spike matrix of the i-th sample in S_{in}
    for $t = 1$ to T **do**
      find $firingneuron$ indices $\tau = \{firingneurons \text{ in } \iota\} \cup \{k \in \chi \setminus \iota, P_k > \beta\}$
      for all $j \in \tau$ **do**
        find post synaptic neuron indices γ
        for all $k \in \gamma$ and $R_k = 0$ **do**
          $P_k \leftarrow P_k + w_{jk}$ {update potential}
        end for
      end for
      $P_k \leftarrow 0, \forall k \in \tau$ {reset potential}
      $R_k \leftarrow R, \forall k \in \tau$ {reset refractory counter}
      $P_k \leftarrow max(0, P_k - D), \forall k \in \chi \setminus \iota \setminus \tau$
      $R_k \leftarrow max(0, R_k - 1), \forall k \in \chi \setminus \iota \setminus \tau$
      for all $j \in \tau$ **do**
        find post synaptic neuron indices γ
        for all $k \in \gamma$ **do**
          $w_{jk} \leftarrow w_{jk} - \eta(t - t_k^f)$
        end for
        find pre synaptic neuron indices γ
        for all $k \in \gamma$ **do**
          $w_{jk} \leftarrow w_{jk} + \eta(t - t_k^f)$
        end for
      end for
    end for
  end for
end for

From the deSNN procedure, described in Algorithm 4, it can be seen that the static output representation of the Spatio-temporal SNNr-based spiking patterns highly depends on the selection of the method's hyperparameters α_{deSNN} and d.

5. Experiments

In [37] it has been suggested to repeat the processing steps described in sections 4.1.1, 4.2, and 4.2.2 for different hyperparameter values in order to optimise the final classification performance in the supervised learning step. Taking into account the entire proposed methodology, this includes the following hyperparameters: α_{TR} for spike encoding; SWR, D, R, η, β and N_{iter} for STDP-based unsupervised learning of the SNNr; α_{deSNN} and d for the deSNN representation; and finally α_{FS}, n_{FS}, C and σ for the supervised learning step, which includes feature selection and an RBF-SVM classifier. This makes a total of 13 hyperparameters that should be simultaneously optimised in an exhaustive procedure where each iteration includes data preprocessing, STDP-based learning as well as cross-validation of the obtained output representation. However, STDP-based training of SNNs is computationally very expensive and accounts for the most significant proportion of the total execution time.

Algorithm 4: Output deSNN representation: f_{deSNN}

Require: $W \in \mathbb{R}^{N_r \times N_r}$, $s_{in} \in \{-1, 0, 1\}^{T \times N_{input}}$
$\{hyperparameters := \alpha, d\}$
Ensure: $W_{deSNN} \in \mathbb{R}^{1 \times N_r}$
$W_{deSNN} \leftarrow 0^{1 \times N_r}$ {initialize representation to 0}
$F \leftarrow 0^{1 \times N_r}$ {initialize the neuron firing flags}
$c \leftarrow 0$ {initialize the neuron firing order counter}
$S_{cube} \leftarrow$ propagate s_{in} through the SNN defined by W {S_{cube} is a sparse $N_r \times T$ matrix of all neuron firings for the entire sample duration length T}
for $i = 1$ to T **do**
  for $j = 1$ to N_r **do**
    if $S_{cube}(i, j) = 1$ **then**
      if $F(j) = 0$ **then**
        {neuron fires for the first time}
        $W_{deSNN}(j) \leftarrow \alpha^c$
        $F(j) \leftarrow 1$
      else
        $W_{deSNN}(j) \leftarrow W_{deSNN}(j) + d$
      end if
      $c \leftarrow c + 1$
    else
      $W_{deSNN}(j) \leftarrow W_{deSNN}(j) - d$
    end if
  end for
end for

In order to mitigate the computationally expensive STDP-based step, we optimise the encoding hyperparameter α_{TR} via a compromise approach described in section 4.1.1, outside of the SNN context, while still taking into account the classification task performance. Parameter $\alpha_{TR} = 1.5$ is identified as optimal for both the arousal and valence classification tasks. For the STDP-based training of the SNNr module, we use a set of previously identified hyperparameters [40]: $SWR = 2.5$, $D = 0.002$, $R = 6$, $\eta = .001$, and $\beta = 0.5$, in order to avoid training the SNNr at each step of the deSNN and supervised learning hyperparameter optimisation. N_{iter} was set to 5, due to the relatively small amount of available EEG data. This parameter has to be taken with caution as high values may cause over-training of the SNNr [37]. The resulting brain-like SNNr connectivity can be visualised, analysed and interpreted for a better understanding of the EEG data and relative involvement of various brain regions. Figure 8 shows the emerging SNNr connectivity, obtained by using data with different emotional labels.

For each of the 4 available subtasks, i.e. arousal and valence classification with DEAP and MAHNOB-HCI, a Bayesian optimisation approach is utilised in order to identify the optimal set of the 6 remaining hyperparameters. α_{FS}, n_{FS}, C and σ are optimised using the same ranges as in section 4.1.1. The range for α_{deSNN} was set to $[10^{-2}, 2]$, and range for d was set to $[10^{-3}, 1]$. At each iteration both simple spike-based features from section 4.1.1, as well as the deSNN-based features, were fed to the feature selection and LOSO cross-validated RBF-SVM evaluation, in search of the maximum accuracy. The idea was to test the added value of deSNN-based features in terms of classification accuracy improvement. Results are summed up in Table 1.

The upper half of Table 1 is based on the results obtained in section 4.1.1 (Figure 6), while the lower half sums up the results of deSNN optimisation procedure described in this section. The addition of optimised deSNN representation to the feature set significantly increased the maximal obtainable accuracy in arousal classification, for both datasets. In terms of valence classification, accuracy was slightly improved for DEAP, but not for MAHNOB- HCI.

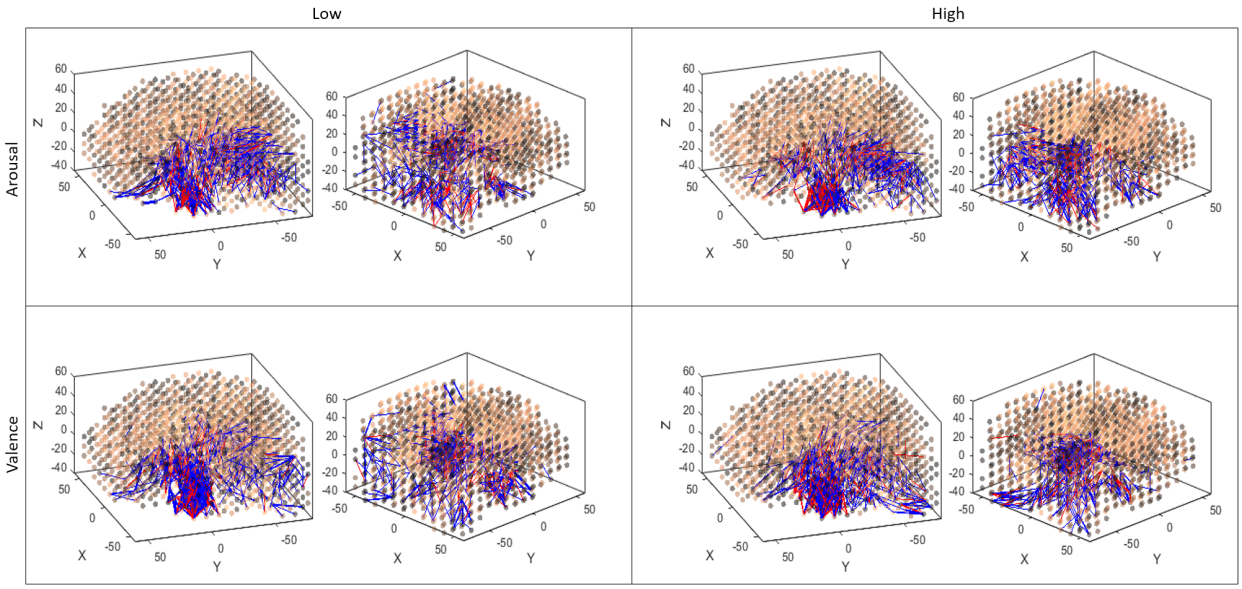

Figure 8: Connectivity of 4 different SNNr modules. Each SNNr is obtained by using the combined data from both DEAP and MAHNOB-HCI, labeled as either low or high in terms of either arousal or valence. For each SNNr two 3-D plots from different angles are given: the left plot shows frontal and left side of the brain, while the right plot shows the back (posterior) and right side of the brain. Neurons are plotted with a slight transparency in order to better highlight the 3D nature of the emerged connections. 500 strongest connections are displayed for each SNNr, with thicker lines denoting stronger connections. Brighter neurons are more active.

6. Discussion

The obtained results demonstrate how complex SNN-based Spatio-temporal modelling of EEG spiking patterns can provide additional information value able for short-term emotion classification. Such SNN-based information was shown to be particularly useful for short-term arousal classification, and not as much for valence.

The main advantage of the described procedure is straight-forward spike-based processing of EEG, as opposed to the need for extraction of handcrafted features in case of the most classical methods, as well as the ability to capture information hidden within the complex Spatio-temporal EEG spiking patterns on relatively short time frames. Furthermore, the addition of subject-specific spike encoding, which is crucial in order to successfully develop subject-independent SNN-based decision systems, as well as the proposed hyperparameter optimisation strategy is an added value to the original NeuCube framework. Hopefully, this will open up additional paths and opportunities for further utilisation of brain-like SNN models in various subject-independent classification tasks.

Table 1: Comparison of the obtained optimized LOSO cross-validation accuracies.

Feature set	Task	Dataset	Accuracy
Simple spike-bassed features	Arousal	DEAP	0.6384
		MAHNOB-HCI	0.7593
	Valence	DEAP	0.6473
		MAHNOB-HCI	**0.7212**
Simple spike-bassed features + deSNN	Arousal	DEAP	**0.7897**
		MAHNOB-HCI	**0.7939**
	Valence	DEAP	**0.6776**
		MAHNOB-HCI	0.7068

From the spike-only-based "comparison approach" (Table 1) emerge both the strengths and limitations of this study. This approach demonstrates that the addition of deSNN features, which are a representation of Spatio-temporal SNNr spiking patterns, to simple non-SNN-related spike features improves the classification performance. In contrast to traditional approaches, an SNN-based emotion classification framework presented in this paper can help in better understanding of short-term emotional processing. From Figure 8, it can be seen that the left side of the brain dominates in terms of connectivity strength, i.e. activation, in particular regions corresponding to the temporal lobe, amygdala and the occipital lobe. This result is not surprising, since both DEAP and MAHNOB-HCI induce emotions by means of emotional videos, which requires visual processing (occipital lobe) as well as the processing of language (temporal lobe). Additionally, it is well known that the amygdala, located in the medial temporal lobe, plays one of the key roles in general emotional processing [77]. The dominance of the left side most likely emerges from the dominance of the left temporal lobe in most people. For example, the lateral sulcus, area considered to be highly involved in language function, is longer on the left than on the right side of the brain [78]. Additionally, the posterior views on the obtained connectivities from Figure 8 suggest that occipital lobe plays the most significant role in the processing of high arousal. This might indicate that the visual portion of the affective stimuli plays the most significant role in inducing high emotional arousal. Due to the indirect presence of the EEG data from all subjects in both the training and validation, as a result of a single unsupervised SNNr training for each of the 4 subtasks, the reported accuracies likely overestimate the true predictive power of framework, and as such call for future work, with bigger amounts of data. Given that, as well as the hyperparameter optimisation strategy described in section 5, Table 1 should rather serve as a demonstration of "added value" in comparison to the classical non-SNN- related spike-based features. Additionally, analysis of the deSNN features in the context of the traditional time- or frequency-domain EEG features would be a useful but exceeds the scope of our work as such features do not match the idea of biologically inspired spike-based processing of EEG. However, the proposed approach should not be considered only in terms of the obtained classification performance, which is the case in most traditional "black box" emotion recognition models. One of the main advantages of this approach is in the understanding of the induced and recognised emotions through a brain-inspired SNN model, as shown in Figure 8 and discussed above.

7. Conclusion

To conclude, the paper for the first time introduces a brain-inspired SNN architecture to recognise, and most importantly, to explain for a better understanding, EEG data measuring two basic emotions: arousal and valence. The patterns obtained after deep learning in the SNN architecture are interpreted in terms of brain activities of subjects. The proposed method can manifest fast, incremental and transfer learning on new data related to emotions that make it suitable for further study and for real-time emotion recognition systems. Future work will involve improved SNN hyperparameter optimisation and better visualisation of the brain- structured SNN during and after learning for a better understanding of brain processes related to emotional processing in humans.

References

[1] R. W. Picard, Affective computing-mit media laboratory perceptual computing section technical report no. 321, Cambridge, MA 2139 (1995).
[2] J. Tao, T. Tan, Affective computing: A review, in: International Conference on Affective computing and intelligent interaction, Springer, 2005, pp. 981–995.
[3] T. J. Brigham, Merging technology and emotions: Introduction to affective computing, Medical reference services quarterly 36 (4) (2017) 399–407.
[4] B. Guthier, R. Dörner, H. P. Martinez, Affective computing in games, in: Entertainment Computing and Serious Games, Springer, 2016, pp. 402–441.
[5] E. Yadegaridehkordi, N. F. B. M. Noor, M. N. B. Ayub, H. B. Affal, N. B. Hussin, Affective computing in education: A systematic review and future research, Computers & Education 142 (2019) 103649.
[6] S. Duo, L. X. Song, An e-learning system based on affective computing, physics Procedia 24 (2012) 1893–1898.
[7] A. Luneski, E. Konstantinidis, P. Bamidis, Affective medicine, Methods of information in medicine 49 (03) (2010) 207–218.
[8] M. S. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, H. Meng, A. Kemp, M. Shafizadeh, et al., The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal emopain dataset, IEEE transactions on affective computing 7 (4) (2015) 435–451.
[9] R. W. Picard, J. Scheirer, The galvactivator: A glove that senses and communicates skin conductivity, in: Proceedings 9th Int. Conf. on HCI, 2001.

[10] F. Cid, J. Moreno, P. Bustos, P. Núnez, Muecas: a multi-sensor robotic head for affective human robot interaction and imitation, Sensors 14 (5) (2014) 7711–7737.
[11] J. Grafsgaard, J. B. Wiggins, K. E. Boyer, E. N. Wiebe, J. Lester, Automatically recognizing facial expression: Predicting engagement and frustration, in: Educational Data Mining 2013, 2013.
[12] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, Deap: A database for emotion analysis; using physiological signals, IEEE Transactions on Affective Computing 3 (1) (2012) 18–31.
[13] M. Soleymani, J. Lichtenauer, T. Pun, M. Pantic, A multimodal database for affect recognition and implicit tagging, IEEE Transactions on Affective Computing 3 (1) (2012) 42–55.
[14] B. Ko, A brief review of facial emotion recognition based on visual information, sensors 18 (2) (2018) 401.
[15] B. Schuller, G. Rigoll, M. Lang, Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture, in: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1, IEEE, 2004, pp. I–577.
[16] I. Mijić, M. Šarlija, D. Petrinović, Mmod-cog: A database for multimodal cognitive load classification, in: 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE, 2019, pp. 15–20.
[17] T.-Y. Huang, J.-L. Li, C.-M. Chang, C.-C. Lee, A dual-complementary acoustic embedding network learned from raw waveform for speech emotion recognition, in: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE, 2019, pp. 83–88.
[18] S. Piana, A. Stagliano, F. Odone, A. Verri, A. Camurri, Real-time automatic emotion recognition from body gestures, arXiv preprint arXiv:1402.5047 (2014).
[19] A. Camurri, I. Lagerlöf, G. Volpe, Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques, International journal of human-computer studies 59 (1-2) (2003) 213–225.
[20] D. Kukolja, S. Popović, M. Horvat, B. Kovač, K. Ćosić, Comparative analysis of emotion estimation methods based on physiological measurements for real-time applications, International journal of human-computer studies 72 (10-11) (2014) 717–727.
[21] A. Greco, G. Valenza, L. Citi, E. P. Scilingo, Arousal and valence recognition of affective sounds based on electrodermal activity, IEEE Sensors Journal 17 (3) (2016) 716–725.
[22] X. Zhang, C. Xu, W. Xue, J. Hu, Y. He, M. Gao, Emotion recognition based on multichannel physiological signals with comprehensive nonlinear processing, Sensors 18 (11) (2018) 3886.
[23] P. Rani, C. Liu, N. Sarkar, E. Vanman, An empirical study of machine learning techniques for affect recognition in human–robot interaction, Pattern Analysis and Applications 9 (1) (2006) 58–69.
[24] O. O. Rudovic, Machine learning for affective computing and its applications to automated measurement of human facial affect, in: 2016 International Symposium on Micro-NanoMechatronics and Human Science (MHS), IEEE, 2016, pp. 1–1.
[25] S. M. Alarcao, M. J. Fonseca, Emotions recognition using eeg signals: A survey, IEEE Transactions on Affective Computing (2017).
[26] C. Tan, M. Šarlija, N. Kasabov, Spiking neural networks: Background, recent development and the neucube architecture, Neural Processing Letters (2020) 1–27.
[27] N. K. Kasabov, Evolving spiking neural networks, in: Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence, Springer, 2018, Ch. 5, pp. 169–199.
[28] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436–444.
[29] W. Wang, G. Pedretti, V. Milo, R. Carboni, A. Calderoni, N. Ramaswamy, A. S. Spinelli, D. Ielmini, Learning of spatiotemporal patterns in a spiking neural network with resistive switching synapses, Science advances 4 (9) (2018) eaat4752.
[30] A. Taherkhani, A. Belatreche, Y. Li, G. Cosma, L. P. Maguire, T. M. McGinnity, A review of learning in biologically plausible spiking neural networks, Neural Networks 122 (2020) 253–272.
[31] N. Kasabov, K. Dhoble, N. Nuntalid, G. Indiveri, Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition, Neural Networks 41 (2013) 188–201.
[32] W. Maass, Networks of spiking neurons: the third generation of neural network models, Neural networks 10 (9) (1997) 1659–1671.
[33] W. Maass, H. Markram, On the computational power of circuits of spiking neurons, Journal of computer and system sciences 69 (4) (2004) 593–616.
[34] S. M. Bohte, J. N. Kok, H. La Poutre, Error-backpropagation in temporally encoded networks of spiking neurons, Neurocomputing 48 (1-4) (2002) 17–37.
[35] S. M. Bohte, H. La Poutré, J. N. Kok, Unsupervised clustering with spiking neurons by sparse temporal coding and multilayer rbf networks, IEEE Transactions on neural networks 13 (2) (2002) 426–435.
[36] B. Meftah, O. Lezoray, A. Benyettou, Segmentation and edge detection based on spiking neural network model, Neural Processing Letters 32 (2) (2010) 131–146.
[37] N. Kasabov, E. Capecci, Spiking neural network methodology for modelling, classification and understanding of eeg spatio-temporal data measuring cognitive processes, Information Sciences 294 (2015) 565–575.
[38] N. Kasabov, J. Hu, Y. Chen, N. Scott, Y. Turkova, Spatio-temporal eeg data classification in the neucube 3d snn environment: methodology and examples, in: International Conference on Neural Information Processing, Springer, 2013, pp. 63–69.
[39] K. Kumarasinghe, N. Kasabov, D. Taylor, Deep learning and deep knowledge representation in spiking neural networks for brain-computer interfaces, Neural Networks 121 (2020) 169–185.
[40] E. Capecci, Z. G. Doborjeh, N. Mammone, F. La Foresta, F. C. Morabito, N. Kasabov, Longitudinal study of alzheimer's disease degeneration through eeg data analysis with a neucube spiking neural network model, in: 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, 2016, pp. 1360–1366.
[41] S. Ghosh-Dastidar, H. Adeli, Improved spiking neural networks for eeg classification and epilepsy and seizure detection, Integrated Computer-Aided Engineering 14 (3) (2007) 187–212.
[42] Z. G. Doborjeh, M. Doborjeh, N. Kasabov, Eeg pattern recognition using brain-inspired spiking neural networks for modelling human decision processes, in: 2018 International Joint Conference on Neural Networks (IJCNN), IEEE, 2018, pp. 1–7.

[43] D. Taylor, N. Scott, N. Kasabov, E. Capecci, E. Tu, N. Saywell, Y. Chen, J. Hu, Z.-G. Hou, Feasibility of neucube snn architecture for detecting motor execution and motor intention for use in bciapplications, in: Neural Networks (IJCNN), 2014 International Joint Conference on, IEEE, 2014, pp. 3221–3225.
[44] J. Hu, Z.-G. Hou, Y.-X. Chen, N. Kasabov, N. Scott, Eeg-based classification of upper-limb adl using snn for active robotic rehabilitation, in: Biomedical Robotics and Biomechatronics (2014 5th IEEE RAS & EMBS International Conference on, IEEE, 2014, pp. 409–414.
[45] Z. G. Doborjeh, N. Kasabov, M. G. Doborjeh, A. Sumich, Modelling peri-perceptual brain processes in a deep learning spiking neural network architecture, Scientific reports 8 (1) (2018) 1–13.
[46] Z. Doborjeh, M. Doborjeh, T. Taylor, N. Kasabov, G. Y. Wang, R. Siegert, A. Sumich, Spiking neural network modelling approach reveals how mindfulness training rewires the brain, Scientific reports 9 (1) (2019) 1–15.
[47] S. Thorpe, J. Gautrais, Rank order coding, in: Computational neuroscience, Springer, 1998, pp. 113–118.
[48] N. K. Kasabov, Evolving connectionist systems: the knowledge engineering approach, Springer Science & Business Media, 2007.
[49] S. G. Wysoski, L. Benuskova, N. Kasabov, Evolving spiking neural networks for audiovisual information processing, Neural Networks 23 (7)(2010) 819–835.
[50] N. K. Kasabov, Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence, Springer, 2018.
[51] N. K. Kasabov, NeuCube: A Spiking Neural Network Architecture for Mapping, Learning and Understanding of Spatio-Temporal Brain Data, Neural Networks vol.52 (2014), pp. 62-76, http://dx.doi.org/10.1016/j.neunet.2014.01.006
[52] N. Kasabov, N. M. Scott, E. Tu, S. Marks, N. Sengupta, E. Capecci, M. Othman, M. G. Doborjeh, N. Murli, R. Hartono, et al., Evolving spatio-temporal data machines based on the neucube neuromorphic framework: design methodology and selected applications, Neural Networks 78 (2016) 1–14.
[53] W.-L. Zheng, B.-L. Lu, Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks, IEEE Transactions on Autonomous Mental Development 7 (3) (2015) 162–175.
[54] S. Katsigiannis, N. Ramzan, Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices, IEEE journal of biomedical and health informatics 22 (1) (2017) 98–107.
[55] P. Viola, M. Jones, et al., Rapid object detection using a boosted cascade of simple features, CVPR (1) 1 (511-518) (2001) 3.
[56] S. Jianbo, C. Tomasi, Good features to track, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994, pp. 593–600.
[57] B. D. Lucas, T. Kanade, et al., An iterative image registration technique with an application to stereo vision, in: International Joint Conference on Artificial Intelligence, Vancouver, British Columbia, 1981.
[58] V. Kazemi, J. Sullivan, One millisecond face alignment with an ensemble of regression trees, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1867–1874.
[59] M. Soleymani, F. Villaro-Dixon, T. Pun, G. Chanel, Toolbox for emotional feature extraction from physiological signals (teap), Frontiers in ICT 4 (2017) 1.
[60] W.-L. Zheng, J.-Y. Zhu, B.-L. Lu, Identifying stable patterns over time for emotion recognition from eeg, IEEE Transactions on Affective Computing (2017).
[61] Y. Huang, J. Yang, S. Liu, J. Pan, Combining facial expressions and electroencephalography to enhance emotion recognition, Future Internet 11 (5) (2019) 105.
[62] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, J.-H. Chen, Eeg-based emotion recognition in music listening, IEEE Transactions on Biomedical Engineering 57 (7) (2010) 1798–1806.
[63] R. Jenke, A. Peer, M. Buss, Feature extraction and selection for emotion recognition from eeg, IEEE Transactions on Affective Computing 5 (3) (2014) 327–339.
[64] B. Hjorth, Eeg analysis based on time domain properties, Electroencephalography and clinical neurophysiology 29 (3) (1970) 306–310.
[65] S. K. Hadjidimitriou, L. J. Hadjileontiadis, Toward an eeg-based recognition of music liking using time-frequency analysis, IEEE Transactions on Biomedical Engineering 59 (12) (2012) 3498–3510.
[66] B. Petro, N. Kasabov, R. M. Kiss, Selection and optimization of temporal spike encoding methods for spiking neural networks, IEEE transactions on neural networks and learning systems (2019).
[67] S. M. Bohte, The evidence for neural information processing with precise spike-times: A survey, Natural Computing 3 (2) (2004) 195–206.
[68] N. Nuntalid, K. Dhoble, N. Kasabov, Eeg classification with bsa spike encoding algorithm and evolving probabilistic spiking neural network, in: International Conference on Neural Information Processing, Springer, 2011, pp. 451–460.
[69] T. Delbruck, jaer open source project, http://jaer.wiki.sourceforge.net (2007).
[70] P. Lichtsteiner, T. Delbruck, A 64x64 aer logarithmic temporal derivative silicon retina, in: Research in Microelectronics and Electronics, 2005 PhD, Vol. 2, IEEE, 2005, pp. 202–205.
[71] N. Sengupta, N. Kasabov, Spike-time encoding as a data compression technique for pattern recognition of temporal data, Information Sciences 406 (2017) 133–145.
[72] J. Gautrais, S. Thorpe, Rate coding versus temporal order coding: a theoretical approach, Biosystems 48 (1-3) (1998) 57–65.
[73] G. Roffo, S. Melzi, M. Cristani, Infinite feature selection, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 4202–4210.
[74] G. Roffo, Feature selection library (matlab toolbox), arXiv preprint arXiv:1607.01327 (2016).
[75] J. Talairach, P. Tournoux, Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging (1988).
[76] S. Song, K. D. Miller, L. F. Abbott, Competitive hebbian learning through spike-timing-dependent synaptic plasticity, Nature neuroscience 3 (9) (2000) 919.
[77] M. Weymar, L. Schwabe, Amygdala and emotion: the bright side of it, Frontiers in neuroscience 10 (2016) 224.
[78] G. H. Yeni-Komshian, D. A. Benson, Anatomical study of cerebral asymmetry in the temporal lobe of humans, chimpanzees, and rhesus monkeys, Science 192 (4237) (1976) 387–389.